IN THE SPECIFICATION

Please amend the Abstract as follows:

~~Video~~ A system for estimating audience state using the video signal relative to an imaged audience and audio signal according to voices from the audience ~~are~~ generated in an input unit. A characteristic amount detection unit detects information on a movement amount, movement periodicity, a volume, voice periodicity of the audience, and a frequency component of voices from the audience based on the video signal or the audio signal. An estimation unit estimates an audience state based on the detected result. An output unit outputs the estimated result of the audience state. The audience state can be easily estimated without observing the audience state by a person.